AD

HDL Proj: 01221

HDL-TM-70-25

COMPARISON OF HDL SDI SERVICES BASED ON A PRELIMINARY IEEE TAPE AND ON DDC TAPES

by

Berthold Altmann

October 1970

U.S. ARMY MATERIEL COMMAND

HARRY DIAMOND LABORATORIES

WASHINGTON, D.C. 20438

ABSTRACT

The results of the HDL SDI service measured in numbers of items requested by its subscribers are presented and analyzed for each of two types of bulletins: one derived from the DDC tapes on reports literature, and one based on a preliminary IEEE tape containing open (periodical) literature. The latter contains the information of only one-half the August 1970 issues.

For more than one year ending in July 1970, the Technical Information Office of HDL processed the tape records furnished without charge by DDC to obtain information related to the installation's missions and tasks. The tapes contained the bibliographic contents of the Technical Announcement Bulletins issued by DDC and of the Research and Development Reports Clearinghouse Abstract Journal. Items were extracted on the basis of pertinent COSATI fields and groups, and the resulting tape and its selected data provided an SDI (Selective Dissemination of Information) service for 45 teams whose key personnel had submitted accurately formulated requirements. This service and its principles and methodology were described briefly in a recent report.[1]

After several months of operation it became apparent that this current awareness program that we had initiated was appreciated by the subscribers; not only did they express their appreciation, but also they responded by consistently and promptly returning lists with requests for acquisitions and loans. Moreover, in many instances, the subscribers visited our analysts to assist in providing more precise formulations of their individual requirements. It was especially pleasing that the majority of the personnel served were the engineers who previously had not utilized the collection on a regular basis.

The DDC tape service has greatly helped to render a personalized service to our laboratory personnel without placing an intolerable strain on the limited resources of the installation. However, it was obvious that in order to become fully adequate, this reference service had to be extended to encompass the rapidly growing periodical literature. The tapes that the IEE in London produced in conjunction with the publication of Science Abstracts, which the IEEE in the United States adjusted and augmented to meet in particular the requirements of American electronics engineers, seemed to offer the most appropriate and economical means of achieving our objectives, to extend the coverage of the SDI service as well as to build a broader base for future computerized bibliographic and on-line retrieval operations.

The installation has subscribed (a) to all chapters of the Electrical and Electronic Abstracts except chapter (4), Power and Industry, and (b) to Chapter 16.00 and 17.00 of Physics Abstracts, which contain the information on solid-state structure and mechanical properties, and solid-state electrical and magnetic properties.

From the IEEE we received a preliminary tape containing half the information that will be published in the August 1970 issues of Science Abstracts. After minor adjustments of the computer programs that we use in connection with the DDC materials, we purged this tape and generated

---

1. B. Altmann, The HDL Automated Information System, HDL-TR-1523, Harry Diamond Laboratories, Washington, D. C. August 1970

SDI bulletins for our subscribers. In the following pages we present the results compared with those obtained in the SDI service based on the tapes that DDC had distributed. When evaluating the two services which we have not yet integrated, one must keep in mind that the statistics of the DDC-based service represent average results from six bulletins (with an input of 10,765 documents and a purged record of 5,724 documents), whereas the one IEEE tape contained information on only 2,131 documents and yielded a purged tape of 2,105 titles for the selection by profiles.

In determining utility, we were guided by the consideration that it is not the objective of a reference service to provide information whose relationship to a user's requirement or profile is more or less remote. The matching of a number of terms will always assure such results. However, the reference service's overriding purpose is to furnish the papers and documents that are useful and that assist the requestor substantially. For this reason, the evaluation of the service was left entirely to the discretion of the subscriber. What was checked as worthwhile information that should be acquired for the library or for loan was considered a hit, and what was not selected for whatever reason (it could have been an article in a periodical to which he subscribed or one of his own publications) was recorded in the negative column.

The retention of a larger number of data on the purged IEEE tape is due to the effective preselection from the more appropriate groups in Science Abstracts. The relatively smaller usefulness of the purged IEEE tape for the HDL profiles can be explained by the great number of theoretical studies and of studies related to applications for which HDL personnel have not expressed an interest at least up to this time. Nevertheless, the nearly 70 percent greater selection by the subscriber from the "IEEE-service" leads to the conclusion that the profiles are well chosen. Moreover, it is conceivable that in the testing process (of the IEEE tape) the subscribers clearly decided what they considered acceptable and what

Table A

Comparison of SDI Service Based on DDC Tapes 10-15 (1970) and on the test Tape Provide by IEEE (One Half of the August 1970 Issues) (in Percentages)

	Titles retained after purge of HDL-unrelated groups	Unique titles announced in the SDI Bulletins (versus titles on purged tapes)	All titles announced is SDI Bulletins (versus titles on purged tapes)	Selection by subscribers from distributed bulletins	Rejection by subscribers from distributed bulletins
DDC	53	34.6	53.35	23.95	47.05
IEEE	96.3	17	21.00	39.3	60.07

was not, while in the more numerous SDI Bulletins announcing the DDC reports literature they left 29 percent of the titles unmarked, so that the de facto discrepancy might be much smaller than the absolute percentages indicate. We organized and operated the retrieval from the DDC tapes on the basis of the entire vocabulary contained in the titles of the reports and in the descriptors and identifiers that the analysts and indexers in DDC had assigned.

For the IEEE tapes we utilized a slightly different vocabulary that is derived not only from the titles, but also from standardized subject headings as well as from free-style descriptors. Because the free-style descriptors are mainly repetitious and require a relatively large number of bits, they extend time and cost of the retrieval runs considerably. For a sample of 20 randomly selected titles, it was observed that the average number of retrieval terms per title is 5.7. The assigned standard subject headings consist of an average of 7.9 terms of which 3.6 (or 45.6 percent) are supplementary to the title information. The free-style descriptors yield 26.3 terms per title, but only 2.1 (or 8 percent) of these differ from the title and subject heading vocabulary and can therefore improve the retrieval results.

In order to verify the relative ineffectiveness of the probably quite expensive analytical effort required to generate the free-style descriptors we ran our profiles against the combination of title and subject terms (of the 2052 items remaining on the purged IEEE tape). This reduction of the retrieval operation lowered the scope of the SDI lists by 20 percent (or 53 titles), while the subscribers lost only 5.8 percent of the titles they had selected from their SDI bulletins.

Despite the cost of a rather small return, we must retain the free descriptors in the retrieval process, because the recipients of the service would otherwise be deprived of a few very important publications. On the other hand, the editors of the service must recognize that the analytical operation in its present form has not reached the feasible level of efficiency. It could be improved by the introduction of SOP's designed to enhance the scope and consistency of the analysis. An analysis in depth would also identify the nature of the study (e.g. whether it is theoretical, experimental, or concerned with application) and also make known the functional relationships of the various subjects the author has covered. The refined precision and selectivity will become a necessity whenever the analytical data are accumulated to furnish on-line retrieval as well as bibliographic service.

This test will be repeated after 2-3 months when the analysis can be based on a larger, more representative sample of the IEEE tape service.

During the past 3 weeks the number of subscribers to the HDL SDI service has risen from 45 to 62 (i.e. or by 37.8 percent).

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Harry Diamond Laboratories Washington, D. C. 20438	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE
Comparison of HDL SDI Service Based on a Preliminary IEEE Tape and DDC Tapes

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*
Berthold Altmann

6. REPORT DATE October 1970	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS 1
8a. CONTRACT OR GRANT NO. b. PROJECT NO. c. HDL Project: 01221 d.	9a. ORIGINATOR'S REPORT NUMBER(S) HDL-TM-70-25 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*	

10. DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY USAMC

13. ABSTRACT

The results of the HDL SDI service measured in numbers of items requested by its subscribers are presented and analyzed for each of two types of bulletins: one derived from the DDC tapes on reports literature, and one based on a preliminary IEEE tape containing open (periodical) literature. The latter contains the information of only one-half the August 1970 issues.

DD FORM 1473 REPLACES DD FORM 1473, 1 JAN 68, WHICH IS OBSOLETE FOR ARMY USE

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Selective dissemination of information-system	8	3				
HDL information system	8	3				
DDC Automatic Magnetic Tape	8	3				
IEEE Automatic Magnetic Tape	8	3				
Coordinate indexing	8	3				
Information systems	8	3				